Lr34 multi-pathogen resistance ABC transporter: molecular analysis of homoeologous and orthologous genes in hexaploid wheat and other grass species

Krattinger, S G ; Lagudah, E S ; Wicker, T ; Risk, J M ; Ashton, A R ; Selter, L L ; Matsumoto, T ; Keller, B

Abstract: The Triticum aestivum (bread wheat) disease resistance gene Lr34 confers durable, race non-specific protection against three fungal pathogens, and has been a highly relevant gene for wheat breeding since the green revolution. Lr34, located on chromosome 7D, encodes an ATP-binding cassette (ABC) transporter. Both wheat cultivars with and without Lr34-based resistance encode a putatively functional protein that differ by only two amino acid polymorphisms. In this study, we focused on the identification and characterization of homoeologous and orthologous Lr34 genes in hexaploid wheat and other grasses. In hexaploid wheat we found an expressed and putatively functional Lr34 homoeolog located on chromosome 4A, designated Lr34-B. Another homoeologous Lr34 copy, located on chromosome 7A, was disrupted by the insertion of repetitive elements. Protein sequences of LR34-B and LR34 were 97% identical. Orthologous Lr34 genes were detected in the genomes of Oryza sativa (rice) and Sorghum bicolor (sorghum). Zea mays (maize), Brachypodium distachyon and Hordeum vulgare (barley) lacked Lr34 orthologs, indicating independent deletion of this particular ABC transporter. Lr34 was part of a gene-rich island on the wheat D genome. We found gene colinearity on the homoeologous A and B genomes of hexaploid wheat, but little microcolinearity in other grasses. The homoeologous LR34-B protein and the orthologs from rice and sorghum have the susceptible haplotype for the two critical polymorphisms distinguishing the LR34 proteins from susceptible and resistant wheat cultivars. We conclude that the particular Lr34-haplotype found in resistant wheat cultivars is unique. It probably resulted from functional gene diversification that occurred after the polyploidization event that was at the origin of cultivated bread wheat.

DOI: https://doi.org/10.1111/j.1365-313X.2010.04430.x

Lr34 multi-pathogen resistance ABC transporter: molecular analysis of homoeologous and orthologous genes in hexaploid wheat and other grass species

Running title: Comparison of orthologous *Lr34* genes in grasses

Simon G Krattinger[1$], Evans S Lagudah[2], Thomas Wicker[1], Joanna M Risk[2], Anthony R Ashton[2], Liselotte L Selter[1], Takashi Matsumoto[3] and Beat Keller[1*]

[1] Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland,

[2] CSIRO Plant Industry, GPO Box 1600, Canberra, ACT, 2601, Australia, and

[3] Plant Genome Research Unit, Division of Genome and Biodiversity Research, National Institute of Agrobiological Sciences, 2-1-2, Kannondai, Tsukuba, Ibaraki 305-8602, Japan

[*]For correspondence (e-mail bkeller@botinst.uzh.ch, phone +41 (0)44 634 82 30 fax +41 (0)44 634 82 04)

[$]Current address: CSIRO Plant Industry, GPO Box 1600, Canberra, ACT, 2601, Australia

Keywords: *Lr34*, gene diversification, comparative analysis, disease resistance, orthologous genes, polyploid wheat.

Word count: Summary 250; Introduction 714; Results 2,937; Discussion 863; Experimental procedures 891; Acknowledgements 63; table and figure legends 765; Total 6,483; References 1,510.

Summary

The wheat disease resistance gene *Lr34* confers durable, race non-specific protection against three fungal pathogens and has been a highly relevant gene for wheat breeding since the green revoluton. *Lr34*, located on chromosome 7D, encodes an ATP-binding cassette (ABC) transporter. Both wheat cultivars with and without *Lr34*-based resistance encode a putatively functional protein that differ by only two amino acid polymorphisms. In this study we focused on the identification and characterization of homoeologous and orthologous *Lr34* genes in hexaploid wheat and other grass species. In hexaploid wheat we found an expressed and putatively functional *Lr34* homoeolog located on chromosome 4A, designated *Lr34-B*. Another homoeologous *Lr34* copy, located on chromosome 7A, was disrupted by the insertion of several repetitive elements. Protein sequences of LR34-B and LR34 were 97% identical. Orthologous *Lr34* genes were detected in the genomes of rice and sorghum. Maize, *Brachypodium distachyon*, and barley were lacking *Lr34* orthologs, indicating independent deletion of this particular ABC transporter gene. The *Lr34* gene was part of a gene-rich island on the wheat D genome. We found gene colinearity on the homoeologous A and B genomes of hexaploid wheat, but little microcolinearity in other grasses. The homoeologous LR34-B protein and the orthologs from rice and sorghum have the susceptible haplotype for the two critical polymorphisms distinguishing the LR34 proteins from susceptible and resistant wheat cultivars. We conclude that the particular haplotype of the *Lr34* allele found in resistant wheat cultivars is unique. It probably resulted from functional gene diversification that occurred after the polyploidization event that was at the origin of cultivated bread wheat.

Introduction

Selection and development of crops with durable and race non-specific disease resistance constitute important breeding objectives. Together with rice and maize, wheat ranks among the three most important crops in the world. Yield losses caused be fungal pathogens are substantial (Kolmer 2005, Strange and Scott 2005). Unfortunately, many resistance genes in the wheat gene pool have been overcome in the field due to rapid adaptation of pathogens. Only a handful of genes have been reported to protect crops over several decades. Recent progress in map-based cloning of those durable disease resistance genes shed light on the underlying proteins and potential molecular mechanisms (reviewed by Poland *et al.* 2009, Kou and Wang 2010). In barley the recessive *mlo*-allele confers broad spectrum resistance to barley powdery mildew (*Erysiphe graminis* f. ssp. *hordei*). *Mlo* encodes a 60 kDa membrane anchored protein with seven membrane-spanning helices (Buschges *et al.* 1997). The rice gene *pi21* confers recessive durable resistance against blast disease (*Magnaporthe oryzae). pi21* encodes a proline rich protein that includes a heavy metal-binding domain (Fukuoka *et al.* 2009). The authors suggested that metal transport by wild-type *Pi21* might be associated with slowing plant defence. The stripe rust resistance gene *Yr36* of wild emmer confers temperature-dependent, race non-specific resistance against *Puccinia striiformis*. It encodes a protein with a kinase and a START domain. START domains are thought to be involved in lipid trafficking (Fu *et al.* 2009). Finally, *Lr34* located on wheat chromosome 7DS in hexaploid wheat, confers durable resistance against leaf rust (*Puccinia triticina*), stripe rust (*P. striiformis*), and powdery mildew (*Blumeria graminis*). Because of its durability and ability to restrict growth of multiple pathogens, *Lr34* has become one of the most important disease resistance genes in wheat breeding.

It has been incorporated into more than 50% of wheat cultivars around the world and was an important component of disease resistance in the development of high-yielding cultivars during the green revolution (Hoisington *et al.* 1999). *Lr34* was found to encode a full-size ATP-binding cassette (ABC) transporter of the ABCG (formerly pleiotropic drug resistance; PDR) sub-family (Krattinger *et al.* 2009). From these studies it became obvious that diverse molecular mechanisms are involved in durable, race non-specific disease resistance. This is in contrast to many race-specific resistance (*R*) genes which often encode proteins of the nucleotide-binding site leucine-rich repeat (NBS-LRR) family (Keller *et al.* 2005). Despite the recent progress in identifying durable disease resistance genes, little is known about how these genes evolved in ancestors of cultivated crops and during human selection.

Full-size, monomeric ABC transporters share a conserved structure consisting of two cytosolic nucleotide binding domains (NBD) and two transmembrane domains (TMD). The TMDs form a translocation pathway, allowing the substrate to cross biological membranes. Transport is energized by MgATP that binds to the highly conserved NBDs (Jasinski *et al.* 2003). Full-size transporters of the ABCG sub-family are thought to have arisen through a single ancestral duplication of a half-size, dimeric ABCG transporter (formerly White Brown Complex) that consisted of only one NBD and one TMD (Crouzet *et al.* 2006). The ABCG subclass of ABC transporters possibly transports a wide set of structurally and functionally diverse molecules (Rea 2007). The genome sequences of Arabidopsis and rice contained 15 and 23 full size ABCG transporters, respectively (Crouzet *et al.* 2006). Hence, we may expect up to 60 full-size ABCG transporter genes in hexaploid wheat with its three homoeologous genomes.

Interestingly, both wheat cultivars with and without *Lr34*-based resistance were found to carry an expressed and putatively functional *Lr34*-gene on chromosome 7DS. The nucleotide sequence of *Lr34* spans 11,805 bp and consist of 24 exons. Alleles of resistant (*Lr34res-D*) and susceptible (*Lr34sus-D*) cultivars differed by only three polymorphisms, two of which resulted in amino acid changes in the ABCG transporter. A deletion of the three base pairs 'TTC' in exon 11 of the resistant allele resulted in the deletion of a phenylalanine residue at position 546 in LR34res-D, and a C/T SNP in exon 12 converted a tyrosine to a histidine at position 634. Germplasm characterization of 700 wheat accessions revealed that most cultivars carried one of the two haplotypes described above (Dakouri *et al.* 2010). More importantly, almost all accessions that were associated with an *Lr34* resistance phenotype showed the deletion of the phenylalanine and had a histidine at position 634. The only two exceptions known today are cultivars 'Jagger' and 'H45' that are classified as susceptible but showed the *Lr34res-D* haplotype. Sequencing of the *Lr34-D* gene in 'Jagger' identified a point mutation that resulted in a premature stop codon (Lagudah *et al.* 2009).

Modern bread wheat is an allohexaploid grass species (AABBDD genome, 2n=6x=42) that arose through hybridization of three related diploid grasses. The combination of several very similar genomes results in gene multiplication and redundancy. Polyploidy increases the potential for genes to evolve new functions (Bottley *et al.* 2006, Pumphrey *et al.* 2009).

In this study we investigated the structure of homoeologous and orthologous copies of the *Lr34-D* gene in hexaploid wheat and other grasses. On the wheat B genome we found a putatively functional *Lr34* homoeolog that was 97% identical to the *Lr34* resistance gene. The *Lr34* copy on the A genome was disrupted by insertion of

repetitive elements. The data obtained suggest that the *Lr34* resistance allele is unique and evolved after the hybridization of tetraploid wheat with the D-genome donor *Ae. tauschii*, the event that resulted in modern, hexaploid bread wheat.

Results

The hexaploid wheat genome contains three homoeologous Lr34 *genes*

To identify the number of *Lr34*-related sequences in the hexaploid wheat genome we probed Southern blots of the hexaploid wheat cultivar 'Chinese Spring' with a fragment spanning exons 10 to 11 of the *Lr34-D* gene. In total, three bands were observed. Using nulli-tetrasomic lines (Sears 1954) they were mapped to the homoeologous chromosomes 7A, 4A (contains a translocated part of chromosome 7BS), and 7D (figure 1, left panel). For the map-based cloning of *Lr34-D* BAC libraries have been searched (Krattinger *et al.* 2009). During that work, we had identified four BAC clones that contained *Lr34*-related sequences which did not map to chromosome 7D. These four clones were probed on a Southern blot containing mixtures of BAC DNA and DNA of nulli-tetrasomic wheat lines. Using this approach, clones ABCT5 and ABCT33 could be assigned to chromosome 4A and clone ABCT16 to chromosome 7A (figure 1, right panel). The band of clone 306D02 was slightly smaller than the expected 7A band. This shorter fragment was due to the truncation of the *Lr34*-related sequence at the BAC end. Thus, despite the high number of full-size ABCG transporters in plant genomes we detected only three *Lr34*-related sequences in the hexaploid wheat cultivar 'Chinese Spring', indicating that there is a single copy of this gene per diploid genome.

We fully sequenced BAC clones ABCT5 and ABCT33 of chromosome 4A as well as ABCT16 and 306D02 of chromosome 7A. While ABCT33 spanned the entire ABCT5 clone, the two clones ABCT16 and 306D02 showed an overlap of 7.2 kb. Clones ABCT5, ABCT33, and ABCT16 were derived from the Canadian wheat cultivar 'Glenlea', whereas clone 306D02 was isolated from 'Chinese Spring'. Although being of different origin, the overlapping regions of ABCT16 and 306D02 showed only 8 SNPs in 7,203 bp, suggesting a very high sequence similarity between the two cultivars. This finding was consistent with the results of Wicker *et al.* (2009) who reported a very high sequence similarity between 'Glenlea' and 'Chinese Spring' in the *Lr34-D* target interval on chromosome 7D (1 SNP in 146,245 aligned bases). For both the 7A and 4A chromosomes, we found a full-length *Lr34*-like gene, referred to as *Lr34-A* and *Lr34-B*, respectively (figure 2). Exon-intron structures of *Lr34-A* and *Lr34-B* were predicted based on the genomic and cDNA sequence of *Lr34-D*. The coding sequence of *Lr34-A* was interrupted by the insertion of several repetitive elements. Three of these elements, a 303 bp hAT transposon in exon 2, a 5.2 kb LTR retrotransposon in exon 10, and a CACTA transposon in exon 19, resulted in disruption of the *Lr34-A* coding sequence. In addition, two nested CACTA elements in intron 16 enlarged the size of the respective intron from 99 bp to more than 16 kb. The 3' end of the gene showed two partial duplications that involved parts of a CACTA transposon and exons 19 to 22 (curly brackets I and II in figure 2). In addition, the gene lacked exons 23 and 24 and the predicted coding sequence terminated at base pair 74 of exon 22. We were not able to fully resolve the duplications because the complexity of the sequences involving the duplicated CACTA element resulted in two sequence gaps that could not be closed. Due to the various transposon insertions, the gene sequence of *Lr34-A* covered more than 40

kb. The second homoeologous copy, *Lr34-B*, possibly encodes a functional protein. Its coding sequence covered 12,592 bp which was 787 bp longer than *Lr34-D*. The main reason for this size difference was the expansion of intron 18 by 743 bp in *Lr34-B* (figure 2). A 230 bp portion of this expansion could be explained by the insertion of two 'Stowaway' miniature inverted-repeat transposable elements (MITES) in intron 18 of *Lr34-B*.

To test for expression of the homoeologous copies, we designed primers on exon sequences that were identical on *Lr34-A*, *Lr34-B*, and *Lr34-D*. The PCR products were cloned into vectors, sequenced, and the sequences matched to one of the homoeologous genes. In the hexaploid Canadian wheat cultivar 'Thatcher *Lr34*' we found expression of *Lr34-B* and *Lr34-D*, but we did not detect transcript of *Lr34-A*. In tetraploid wheat we detected only the *Lr34-B* transcript using cultivar 'Kronos'. These findings are consistent with the disruption of the A genome copy by retroelements, both in tetraploid and hexaploid wheat.

The predicted 4,209 bp mRNA of *Lr34-B* was of exactly the same size as the cDNA sequence of *Lr34sus-D* and the putative LR34-B protein showed 97% amino acid identity to LR34-D. Out of 1,402 amino acids, 36 residues differed between LR34res-D and LR34-B (figure S1). 10 of these polymorphisms were located in the first N-terminal 164 amino acids before NBD1 (1 polymorphism per 16 amino acids). Only 4 amino acid changes were located in the two NBDs (1 polymorphism per 81 amino acids) and 10 residues differed among the two TMDs (1 polymorphism per 50 amino acids). These results confirm the high conservation level of NBDs and TMDs (Crouzet *et al.* 2006, Rea 2007). The N-terminal residues of ABCG transporters have been observed to be very variable, suggesting that this region is under less structural constraint.

Only two exon-polymorphisms in exons 11 and 12 were detected between *Lr34-D* alleles of susceptible and resistant wheat cultivars (Krattinger *et al.* 2009). We sequenced exons 11 and 12 in the putatively functional homoeologous *Lr34-B* gene in 16 wheat cultivars that have been well characterized for the presence or absence of *Lr34*-based resistance. All cultivars showed the susceptible haplotype with no phenylalanine-deletion and a tyrosine at position 634 (Table 1).

Twelve transmembrane helices are predicted for Lr34 *related transporters*

LR34sus-D was originally predicted to contain 10 transmembrane helices by the SOSUI (Hirokawa *et al.* 1998) program, the archetypal 6 in the first membrane domain and 4 in the second membrane domain (Krattinger *et al*. 2009). Using the TMAP program (Persson and Argos 1997) which uses a multiple sequence alignment of homologous ABCG transporters as the basis of its prediction, LR34sus-D was predicted to contain the archetypal 12 transmembrane helices, 6 in each membrane domain (figure S1). The Arabidopsis transporter PEN3/ At1g59870/PDR8 which was originally predicted to contain 13 trans-membrane helices (Stein *et al.* 2006) is also predicted to contain 12 transmembrane helices by the TMAP program.

Lr34-homologous genes in other grass species

We constructed a phylogenetic tree including the nine to twelve most homologous ABCG proteins identified by BLAST (Altschul *et al*. 1997) in the completely sequenced genomes of rice, *Brachypodium distachyon*, sorghum, and maize (figure 3a). The LR34-B and LR34-D proteins of wheat were part of a cluster that contained three ABCGs from rice, three from sorghum, as well as one copy each from maize and Brachypodium

(red branch in figure 3). Three proteins shared very high similarities with LR34, one from rice (OsABCG50, 86% amino acid identity) and two from sorghum (Sb01g016775, 74% amino acid identity and Sb01g016770, 72% amino acid identity). The closest homologs in Brachypodium and maize showed much lower amino acid identities to LR34, with 54% for Bradi4g45397 and 58% for GRMZM2G014282, respectively. Thus, we consider the two sorghum proteins and OsABCG50 as orthologs (red box, figure 3a). We refer to the closest homologs in Brachypodium and maize, as well as to the rice and sorghum genes OsABCG41, OsABCG49, and Sb08g002910 as *Lr34*-like. The two sorghum orthologs most likely arose through duplication as they were located next to each other in the genome and their predicted cDNA sequences were 86% identical. Sb01g016670 might be a pseudogene as exons 4 and 5 were separated by 23 kb due to insertion of repetitive elements into intron 4. The construction of phylogenetic trees, including homologous ESTs from other grass species, revealed two additional sequences that clustered together with the LR34 orthologs (figure 3b); FL713867 of switchgrass (*Panicum virgatum*, 746 bp) and CA075859 of sugarcane (*Saccharum* L. ssp. hybrids, 857 bp). Thus, we consider it likely that sugarcane and switchgrass also contain orthologous *Lr34* genes. To test for the presence of a *Lr34* ortholog in barley we probed a Southern blot containing DNA of 14 barley cultivars and 5 wild barley (*Hordeum vulgare* ssp. *spontaneum*) accessions with the same probe used in figure 1. We could not detect hybridization in any of the wild and cultivated lines, whereas the wheat and rice controls gave a signal (figure S2). It has been reported that probes derived from wheat hybridize to barley in the majority of cases (Van Deynze *et al.* 1998). In addition, the most homologous barley EST identified using the *Lr34-D* nucleotide sequence (HD07C03r, 571 bp) showed higher identities to OsABCG41

(86%) and OsABCG49 (85%) than to the *Lr34*-ortholog OsABCG50 (71%). These results suggest that there is no ortholog in barley. Maize, sorghum, switchgrass, and sugarcane belong to the subfamily of Panicoideae, wheat, barley, and Brachypodium to Pooideae, and rice to Ehrhartoideae. Orthologous *Lr34* genes were absent in maize, barley, and Brachypodium. Hence, the ABCG transporter must have been deleted independently multiple times during the evolution of grasses (figure 4).

We analyzed the orthologs from rice and sorghum for the amino acids that corresponded to the critical residues 546 and 634 in LR34-D. As for LR34-B, the rice ortholog and the two orthologs from sorghum had the susceptible haplotype found in wheat cultivars without *Lr34*-based resistance (figure 5). We consider it therefore likely that the *Lr34sus-D* allele represents a more ancient version and that two subsequent mutations in wheat resulted in *Lr34res-D*. To further test this hypothesis, we aligned 68 full-size ABCG proteins from *Arabidopsis thaliana*, rice, Brachypodium, and maize (figure S3). For amino acid site 546, 59 ABCG proteins carried a nonpolar amino acid (1 alanine, 3 methionine, 5 phenylalanine, 10 valine, 15 leucine, and 25 isoleucine), and 8 proteins had a polar residue (1 glutamine, and 7 aspartic acid). LR34res-D was the only protein that had a deletion of an amino acid at this position. Residue 546 of LR34-D was predicted to be part of transmembrane helix 2 of TMD 1 which explains the high frequency of aliphatic residues. The deletion of any residue within a transmembrane helix has the potential to cause a major shift in the residues exposed. Typically an α-helix is formed with 3.6 residues per turn. Loss of the phenylalanine in LR34res-D may alter the orientation of the subsequent residues in transmembrane helix 2, potentially altering the specificity of the transported substrate. Position 634 was found to be even more conserved than residue 546. All ABCG proteins except LR34res-D carried an

aromatic amino acid, either phenylalanine (60) or tyrosine (7). Transmembrane prediction suggested that this residue is located at the cytosolic end of transmembrane helix 4 of TMD 1 (figure 5). Thus, this amino acid may be important for substrate recognition or binding. The LR34-D protein differed by only two amino acid residues between susceptible and resistant cultivars. These specific amino acids found in LR34res-D form a unique haplotype that was not shared by any other ABCG transporter.

Meta-analysis of resistance QTLs in wheat

LR34-B and LR34res-D shared 97% amino acid identity. This raises the possibility of LR34-B also playing a role in durable, race non-specific disease resistance. We performed a detailed meta-analysis, reviewing 21 studies in wheat that identified QTLs which conferred quantitative resistance against the three wheat rusts and powdery mildew (Table S1). Due to a translocation event, *Lr34-B* was transferred from chromosome 7BS to chromosome 4AL in tetraploid and hexaploid wheat. Chromosome 4AL of hexaploid wheat contains translocated parts of chromosomes 5AL and 7BS. The most distal part of chromosome 4AL shows homology to chromosomes 7AS and 7DS and the middle part aligns with chromosomes 5BL and 5DL. Because of a pericentric inversion, the centromeric region of chromosome 4AL is homologous to chromosomes 4BS and 4DS (Naranjo *et al.* 1987, Devos *et al.* 1995, Mickelson-Young *et al.* 1995). Markers defining breakpoints of the individual segments are indicated in figure S4. *Lr34-D* mapped close to the RFLP probe cdo475 on chromosome 7DS. The same probe also mapped to chromosomes 7AS (*Xcdo475-7A*) and 4AL (*Xcdo475-4A*) (Roder *et al.* 1998, Paillard *et al.* 2003, Song *et al.* 2005). A marker developed on the hAT element

insertion of *Lr34-A* co-segregated with *Xcdo475-7A* in the 'Arina' x 'Forno' population used by Schnurbusch *et al*. (2004) (data not shown). We therefore expect *Lr34-B* to be in close association with *Xcdo475-4A*. QTLs on chromosome 7D that mapped to the *Lr34-D* region were reported in 10 of these studies in 8 different wheat cultivars. Three studies detected QTLs on chromosome 4AL (Faris *et al.* 1999, Ramburan *et al.* 2004, Singh *et al.* 2009) including a QTL for leaf rust resistance identified by Faris *et al.* (1999). The QTL peaked around *Xfba211*. This marker mapped proximal of *Xwg622* (Song *et al.* 2005) which defines the 4AS-5AL breakpoint. Hence, this QTL maps close to the centromere and is not part of the 7BS translocation (figure S4). A second QTL for field resistance against stripe rust on chromosome 4AL was reported by Ramburan *et al.* (2004) in the hard, red spring wheat cultivar 'Kariega'. The QTL peaked around locus *Xgwm160* that is found on the 7BS translocation. However, *Xgwm160* mapped 14.4 cM distal to *Xcdo475-4A* in the SSR map of Song *et al.* (2005). It is therefore likely that the QTL reported by Ramburan *et al.* (2004) maps distal to *Lr34-B*. Singh *et al.* (2009) reported a third QTL for leaf rust resistance on chromosome 4A in the European winter wheat cultivar 'Beaver'. The QTL peak was around *Xbarc70* which mapped 49.2 cM distal to *Xcdo475-4A* in the map of Song *et al.* (2005). Hence, we consider it likely that the three QTLs reported on chromosome 4A map to different positions than *Lr34-B*. The LR34-B transporter may therefore have a function different from durable disease resistance or there is no functional variation of this gene in the wheat lines tested.

Comparative genomics of the three homoeologous Lr34- *regions in wheat*

The *Lr34-D* gene was part of a 207 kb gene-rich island on chromosome 7D that, beside *Lr34-D*, contained six additional open reading frames encoding proteins with homology

to a hexose carrier, two cytochrome P450 proteins, two lectin receptor kinases, and a cysteine proteinase (Krattinger *et al.* 2009, figure 6). The two paralogous cytochrome genes were 84% identical, indicating recent gene duplication. This was in contrast to the paralogous kinase genes which only shared 67% identity. We compared this interval to the homoeologous regions on chromosome 7A (BAC clones 306D02 and ABCT16) and chromosome 4A (BAC clones ABCT5 and ABCT33). The sequences of chromosome 7A contained six of the seven genes found in the *Lr34-D* target interval and the assembly of chromosome 4A contained four of the seven open reading frames. Overall, the relative positions of the genes were highly conserved among the three genomes. The only exception was a rearrangement involving one of the two cytochrome P450 coding genes of chromosome 7A that was found upstream of the hexose carrier gene (dark blue box, figure 6a). For the lectin receptor kinase genes of chromosomes 7A and 4A, we found short duplicated fragments that corresponded to the 3' end of the full-length genes. They were located downstream of the respective full-length coding sequences (white boxes in figure 6a). None of the homoeologous sequence stretches contained additional gene insertions, suggesting high conservation of this interval in the three wheat genomes. Intergenic sequences were not conserved between the homoeologous genomes and distances between coding sequences varied considerably. The intergenic distance between the hexose carrier gene and the ABCG transporter gene, for example, measured 27 kb in the A genome (cultivar 'Glenlea'), 33 kb in the B genome (cultivar 'Glenlea'), but only 9 kb in the D genome (cultivar 'Chinese Spring').

The Lr34-interval is not conserved in other grass species

We performed a BLAST search to identify the orthologous coding sequences of the genes found in the *Lr34*-target interval in rice, sorghum, and *Brachypodium*. The interval was not conserved in any of these grass species. The orthologous coding sequences were dispersed over different chromosomes. We found orthologs to the hexose carrier gene located on chromosomes 6, 10, and 1 in rice, sorghum, and Brachypodium, respectively (Figure 6b). The region of rice chromosome 6 and sorghum chromosome 10 are syntenic to wheat group 7 chromosomes (The International Brachypodium Initiative 2010). In Brachypodium, the hexose carrier gene has been transferred to a different chromosomal location. This became obvious when analyzing the neighboring genes which were syntenic between rice and sorghum but different in Brachypodium. The closest homologs of both the wheat ABCG and the cytochrome coding genes were neighboring genes on rice chromosome 12 (OsABCG50 and Os12g32850) and represent the only conserved gene pair in a grass genome other than the wheat homoeologous regions. The second cytochrome gene found in wheat most likely arose through recent duplication as we only found one cytochrome P450 gene in rice and the two wheat sequences were 84% identical. The ABCG transporter gene was duplicated in sorghum. However, there was no cytochrome P450 gene neighboring the two ABCG genes. In Brachypodium, we did not find close orthologs for either the ABCG transporter gene or the cytochrome P450 gene. The orthologous lectin receptor kinase genes were found on rice chromosome 7, sorghum chromosome 2, and chromosome 1 of Brachypodium. According to the International Brachypodium Genome Initiative (2010) these chromosomes are syntenic. The three regions had different numbers of the lectin receptor kinase genes, with two copies found in sorghum, three in rice, and five in Brachypodium.

In summary, we found very little microcolinearity in the studied region among the different grass species. Numerous translocations, duplications, and gene deletions occurred since the divergence of these species around 50 million years ago.

Discussion

In this study we investigated the occurrence and structure of orthologous *Lr34* genes in hexaploid wheat and other grass species. Despite the expected presence of 15 to 23 full-size ABCG transporter genes per diploid wheat genome, Southern hybridization detected only three homoeologous *Lr34*-related sequences in the hexaploid wheat cultivar 'Chinese Spring'. Radiolabeled probes hybridize to genome fragments of at least 80% similarity and they are well suited to detect homologous genes in organisms with incomplete genome sequence (Sorrells *et al.* 2003). Beside the *Lr34-D* resistance gene we found two additional orthologous *Lr34* genes on the homoeologous chromosomes 7A and 4A, named *Lr34-A* and *Lr34-B*, respectively. While the coding sequence of *Lr34-A* was disrupted by five repetitive elements, *Lr34-B* was expressed and encoded a putatively functional protein of 97% identity to LR34res-D. *Lr34-D* was part of a gene rich island that contained six additional coding sequences (Krattinger *et al.* 2009). Comparison of the three homoeologous *Lr34*-intervals revealed gene colinearity among the wheat sub-genomes. The only rearrangement found included one of the two paralogous cytochrome genes on chromosome 7A. Intergenic sequences varied considerably in length and transposon composition. These findings are consistent with other studies in wheat which reported good microcolinearity in genic regions and sequence divergence in intergenic sequences (*Glu-1,* Gu *et al.* 2006 and *Acc,* Chalupska *et al.* 2008). For the homoeologous *Glu-1* regions that control grain quality, the authors

found that microcolinearity among the seven genes was maintained, but intergenic sequences were not conserved. In the homoeologous wheat grain *Hardness* (*Ha*) loci on the other hand, several gene eliminations and rearrangements occurred after the formation of hexaploid wheat (Chantret *et al.* 2005). It has been suggested that synteny levels between homoeologous wheat chromosomes decline along the chromosomes towards the distal, high-recombination ends (Akhunov *et al.* 2003).

The recent completion of whole-genome sequences of rice, sorghum, maize, and Brachypodium (International Rice Genome Sequencing Project 2005, Paterson *et al.* 2009, Schnable *et al.* 2009, The International Brachypodium Initiative 2010) allows detailed comparative analyses between different grass species. This is of particular interest for studying gene number and sequence relatedness among multi-gene families such as ABCG transporters. Orthologous *Lr34* genes were found in the genomic sequences of rice and sorghum. Furthermore, we have evidence of additional *Lr34* orthologs in sugarcane and switchgrass. The absence of close *Lr34* orthologs in maize, Brachypodium, and barley indicates multiple, independent deletions of these ABCG transporter genes. Absence of an *Lr34* orthologous gene in these grass species suggests that the ABCG transporter is not essential for survival or that a more distantly related transporter compensates for the loss of the *Lr34* ortholog. So far there are no functional data available for the *Lr34* orthologs from rice and sorghum that might help elucidate the functions or substrates of the encoded proteins. We noted that two of the three orthologs were incorrectly annotated in the respective genome projects, possibly due to the large gene size and the presence of 23 introns.

The seven genes of the wheat *Lr34* interval were not found in syntenic positions of rice, sorghum, and Brachypodium. The only conservation maintained involved the clustering

of the paralogous lectin receptor kinase genes in sorghum, rice, and Brachypodium, and the rice ortholog OsABCG50 and the neighboring cytochrome P450 gene. Although the concept of colinearity between related grass species at the chromosome level is well accepted, several comparative analyses have revealed a breakdown of microcolinearity between related grass species (Sorrells *et al.* 2003, Bossolini *et al.* 2007,).

Many domesticated crops such as wheat, canola, tobacco, and cotton have allopolyploid genomes. The hybridization of closely related genomes results in gene redundancy. It has been shown that polyploidization initiates rapid genome evolution including DNA removal, reactivation of transposable elements, and diversification of duplicated genes (Gu *et al.* 2006, Udall and Wendel 2006, Woodhouse *et al.* 2010). New combinations of transcription factors and epigenetic mechanisms after polyploidization may result in additional variation of the homoeologous genes transcription level compared to their diploid progenitors (Bottley *et al.* 2006, Pumphrey *et al.* 2009). These molecular mechanisms increase variation in sequence and transcript abundance after polyploidization which may result in inactivation of homoeologous gene copies or acquisition of new functions. In wheat we found two homoeologous LR34 proteins that shared 97% identity. LR34-B and the orthologous proteins from rice and sorghum all showed the susceptible haplotype for the critical amino acid polymorphisms that distinguished LR34sus-D and LR34res-D. Further, 50 accessions of *Aegilops tauschii*, the diploid donor of the wheat D genome, all carried the *Lr34sus-D* haplotype (Kolmer *et al.* 2008). Therefore, the *Lr34sus-D* haplotype is very likely the ancestral allele and the *Lr34res-D* allele may be the result of gene diversification after formation of hexaploid wheat approximately 8,000 years ago. In conclusion, the detailed analysis of *Lr34* gene homologs in the grass family presented in this study has revealed that the two

amino acids that are critical for the function of *Lr34res-D* as a durable resistance gene are evolutionary very young. Based on this finding, we suggest that the two mutations in the ancestral *Lr34sus-D* allele have resulted in a new allele with a novel function; the durable protection against biotrophic pathogens, while the original function of *Lr34sus-D* may still be maintained by *Lr34-B* in resistant wheat cultivars. On a functional level these introduced amino acid changes may affect conformational structure, binding specificity, or protein stability. We therefore speculate that the durable *Lr34*-resistance is the result of gene diversification after hybridization of hexaploid wheat. Thus, human selection may have played an important role in fixing those two unique, spontaneous mutations in wheat germplasm.

Experimental procedures

Southern Blot

For Southern Blot 15 µg of genomic wheat DNA were digested for 3-4 hours with 40 units of *Eco*RI (New England BioLabs). For the digestion of BAC clones we used 0.5 ng per clone. BAC DNA was mixed with genomic DNA of nulli-tetrasomic wheat lines (Sears 1954) to avoid over-digestion of BAC clones. Digested DNA was blotted on membranes (Hybond^{TM}-XL, GE Healthcare). ^{32}P-labelling was done at 65°C according to the standard protocol (Sambrook and Russel 2001) using the NEBlot® kit (New England BioLabs). Membranes were washed with a 0.5x SSC, 0.1% SDS solution at 65°C and exposed to hypersensitive X-ray films (BioMax MS Films, Kodak) for two days. Filters were hybridized with a probe that spanned 372 bp of exons 10 to 11 of the *Lr34-D* gene (base pairs 4,992 – 5,363 of the *Lr34res-D* gene sequence, accession

number FJ436983, primer forward: 5'-gtt caa agc ctg tgg agc aa-3', primer reverse: 5'-gct ggt att gca tat gcc -3').

Sequencing of BAC clones

Clones ABCT5, ABCT16, and ABCT33 were identified from a BAC library of the hexaploid wheat cultivar 'Glenlea' (Nilmalgoda *et al*. 2003). BAC 306D02 was derived from the library of the hexaploid Chinese landrace 'Chinese Spring' (Allouis *et al.* 2003). Screening of these BAC libraries was done as described by Krattinger *et al.* (2009). All BAC clones were fully sequenced with the Sanger method. Shotgun sequences of the two 7A BAC clones assembled in 11 contigs that were interrupted by 10 sequence gaps. For the clones of chromosome 4A, seven contigs were obtained. To determine the relative positions of these contigs we used three different approaches: 1) matching sequences at both sides of the gaps to their respective shotgun clones (average size 4-10kb, 2) comparing target site duplications of repetitive elements that spanned sequence gaps, and 3) taking into account position of vector sequences. Using these approaches we could assemble one scaffold each for chromosome 7A and chromosome 4A that spanned 152.8 kb and 146.2 kb, respectively (accession numbers HM775491 for ABCT16, HM775492 for 306D02, and HM775493 for ABCT5 and ABCT33).

Characterization of homoeologous Lr34 *genes*

Exon-intron structures of the homoeologous *Lr34*-genes were predicted based on the cDNA and genomic sequence of the cloned disease resistance gene *Lr34-D* (Krattinger, *et al.* 2009). Insertions of repetitive elements were predicted based on the Triticeae Repeat Sequence (TREP) database (Wicker *et al.* 2002). Exons 11 and 12 of *Lr34-B* were PCR amplified using *Lr34-B* specific primers (forward primer: 5'-gtg gga ccc gca

aaa gca tac -3', reverse primer: 5'-gcc att ctg gca tgg agg ctg-3', Tm 65°C) and sequences using the ABI® 3730 (Applied Biosystems). To test for expression of the homoeologous *Lr34* genes, we designed primers whose sequences were conserved on *Lr34-A*, *Lr34-B*, and *Lr34-D* (forward primer: 5'-gag tgt ggc tgg gat cat acc-3', reverse primer: 5'-cct gct gtt gac ata agc tc-3', Tm 55°C). Total RNA was extracted from leaves of cultivars 'Thatcher *Lr34*' and 'Kronos' using a TRIzol solution (38% phenol, 0.8M guanidine thiocyanate, 0.4M ammonium thiocyanate, 0.1M sodium acetate pH 5 and 5% glycerol). First-strand cDNA for RT-PCR was synthesized using Superscript II reverse transcriptase (Invitrogen). PCR products were cloned using the StrataClone™ PCR cloning kit (Stratagene) and 30-60 clones per cultivar were sequenced using the M13 primer. Sequences were compared to the genomic sequences of *Lr34-A*, *Lr34-B*, and *Lr34-D*, respectively. To predict the number of transmembrane helices in the LR34 protein, we used the bioinformatic program TMAP (Persson and Argos 1997, http://bioinfo4.limbo.ifm.liu.se/tmap/index.html). This program uses a multiple sequence alignment of homologous proteins to make a more robust prediction of trans-membrane helices than can be obtained from the sequence of a single protein. Twenty ABC transporter sequences were selected from the top hits in a protein BLAST search with LR34sus-D as the query. This captured the sequence diversity in this class of ABC transporters, thus reducing any bias. Included in the prediction was the LR34sus allele and sequences for 14 Arabidopsis (including PEN3/At1g59870), three rice, one grape and a tobacco ABC transporter. The sequences were aligned with ClustalW (Larkin *et al.* 2007). The Clustal alignment was converted to a GCG MSF format at http://www.ebi.ac.uk/Tools/clustalw2/index.html and submitted to the TMAP prediction

program. The output file was edited slightly to make it compatible with TMAP as suggested by Bengt Persson (personal communication).

Phylogenetic analyses

To identify the most homologous protein sequences in rice, *Brachypodium distachyon*, sorghum, and maize we performed BLASTp searches (Altschul *et al.* 1997) against the following databases: 'Genes in MSU Osa1 Rice Pseudomolecules-Protein' of the Rice Genome Annotation Project (Ouyang *et al.* 2007), the Brachypodium 8x release proteins of BrachyBase (The International Brachypodium Initiative 2010), the Sbi1.4 protein models from MIPS/PASA on v1.0 assembly for sorghum (Paterson *et al.* 2009), and the protein coding models from Maizesequence.org release 4a.53 (Schnable *et al.* 2009). Homologous EST sequences were found by blastn against the EST databases available at NCBI. EST nucleotide sequences were subsequently translated into protein sequences. Protein alignments were done using ClustalX with a gap opening penalty of 10 and a gap extension penalty of 0.2 (Larkin *et al.* 2007). Protein sequences of OsABCG50 and Sb01g016770 were re-annotated by hand based on the *Lr34-D* cDNA sequence. Both proteins were found to be wrongly annotated as two genes in the respective databases. To construct the phylogenetic trees, we used PROTPARS of the PHYLIP package with 100 bootstrap replicates (http://evolution.genetics.washington.edu/phylip/).

Acknowledgements

This work was supported by the Swiss National Science Foundation grant 3100A-12706/1, an Advanced Investigator grant of the European Research Council (ERC-2009-AdG 249996, Durableresistance), and a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme (PIOF-GA-2009-252731, Dures). We thank Jorge Dubcovsky for providing seeds of the wheat cultivar 'Kronos'.

Accession numbers: HM775491 (ABCT16), HM775492 (306D02), and HM775493 (ABCT5 and ABCT33).

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Meta QTL analysis including studies that identified QTLs in wheat for quantitative resistance against the three rusts of wheat and powdery mildew.

Figure S1. Alignment of LR34res-D from cultivar 'Chinese Spring' and LR34-B from 'Glenlea'**.**

Figure S2. Southern blot probed with a fragment of the *Lr34* gene**.** The blot contains DNA of 14 barley cultivars and five wild barley accessions.

Figure S3: Alignment of 71 full-size ABCG transporters from rice, Arabidopsis, *Brachypodium*, sorghum, and maize.

Figure S4. Schematic representation of chromosome 4AL indicating the 4A/5AL translocation, the pericentric inverstion, and the 4AL/7BS translocation.

Table 1. Amino acid composition of LR34-B and LR34-D proteins at positions 546 and 634 in a set of wheat cultivars. These were the only two polymorphic residues that differed between resistant (LR34res-D) and susceptible (LR34sus-D) cultivars in LR34-D proteins. In all cultivars the LR34-B protein showed the same haplotype as LR34sus-D.

cultivar	Lr34 based resistance +/-	Amino acid 546: deletion of phenylalanine in LR34-D (exon 11)	Amino acid 546: deletion of phenylalanine in LR34-B (exon 11)	Amino acid 634: Histidine-Tyrosine in LR34-D (exon 12)	Amino acid 634: Histidine-Tyrosine in LR34-B (exon 12)
Chinese Spring	+	Deletion	No deletion	Histidine	Tyrosine
Renan	-	No deletion	No deletion	Tyrosine	Tyrosine
Thatcher Lr34	+	Deletion	No deletion	Histidine	Tyrosine
Thatcher	-	No deletion	No deletion	Tyrosine	Tyrosine
Lalbahadur Lr34	+	Deletion	No deletion	Histidine	Tyrosine
Lalbahadur	-	No deletion	No deletion	Tyrosine	Tyrosine
Jupateco R	+	Deletion	No deletion	Histidine	Tyrosine
Jupateco S	-	No deletion	No deletion	Tyrosine	Tyrosine
Avocet R	+	Deletion	No deletion	Histidine	Tyrosine
Avocet S	-	No deletion	No deletion	Tyrosine	Tyrosine
Forno	+	Deletion	No deletion	Histidine	Tyrosine
Arina	-	No deletion	No deletion	Tyrosine	Tyrosine
Glenlea	+	Deletion	No deletion	Histidine	Tyrosine
Inia 66	-	No deletion	No deletion	Tyrosine	Tyrosine
Frontana	+	Deletion	No deletion	Histidine	Tyrosine
Bezostaja	+	Deletion	No deletion	Histidine	Tyrosine

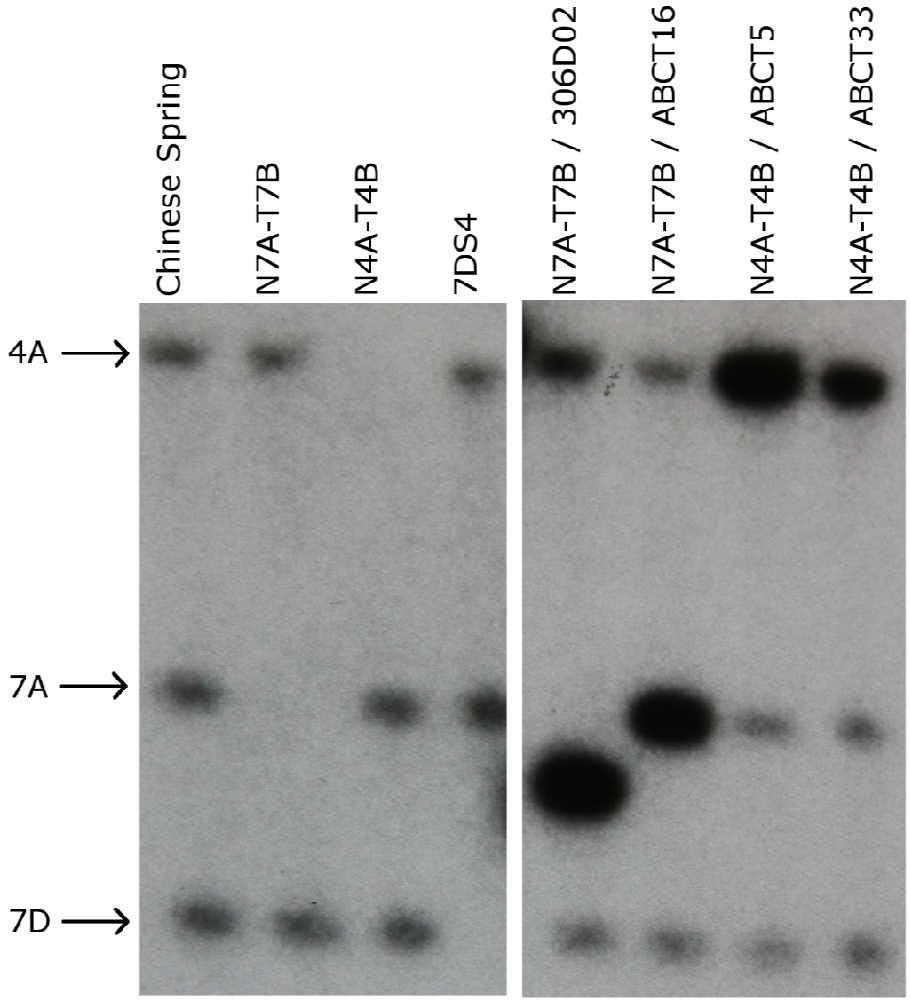


Figure 1. **Southern blot probed with a fragment of the *Lr34-D* gene**. The three bands correspond to the homoeologous chromosomes 7A, 4A, and 7D (left panel). The right panel shows hybridization to mixtures of BAC DNA and DNA of nulli-tetrasomic wheat lines. BAC clones 306D02 and ABCT16 carried a fragment that corresponded to chromosome 7A, whereas ABCT5 and ABCT33 matched to chromosome 4A. N7A-T7B = nullisomic 7A-tetrasomic 7B; N4A-T4B = nullisomic 4A-tetrasomic 4B; 7DS4 = deletion line 7DS4 that lacks *Lr34*. Nulli-tetrasomic and deletion lines were derived from the Chinese landrace 'Chinese Spring'.

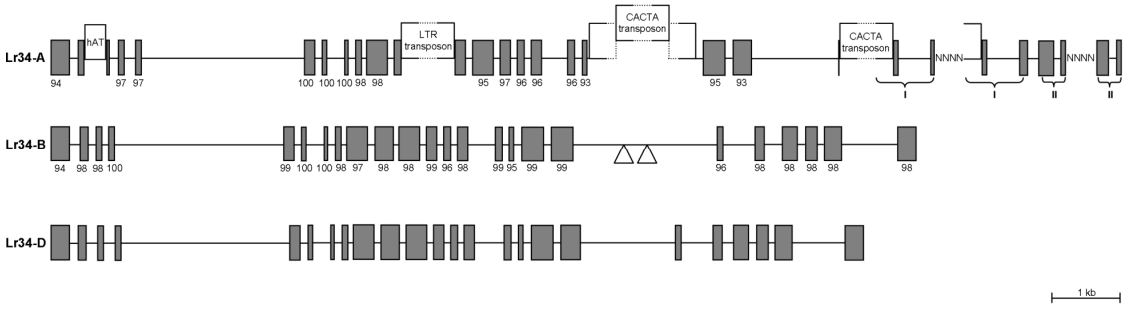
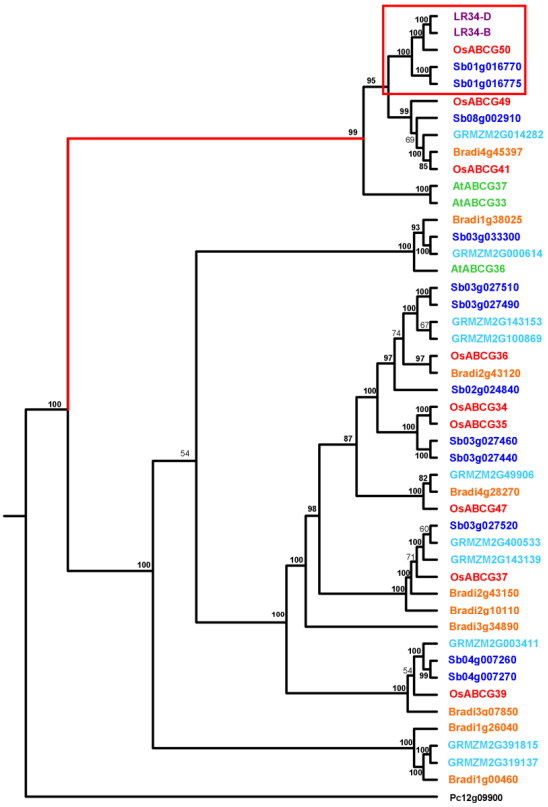


Figure 2. Gene structure of the three homoeologous *Lr34* genes *Lr34-A*, *Lr34-B*, and *Lr34-D*. Grey boxes indicate exons, while introns are shown as adjoining lines. Insertions of repetitive elements in *Lr34-A* (white boxes) are not drawn to scale. Curly brackets delimit duplications at the 3' end of *Lr34-A*. 'N' indicate sequence gaps that could not be closed. Numbers below the exons of *Lr34-A* and *Lr34-B* indicate nucleotide identity of the respective exon to *Lr34-D*. Triangles mark the insertion of two MITE transposable elements in intron 18 of *Lr34-B*.

a)



b)

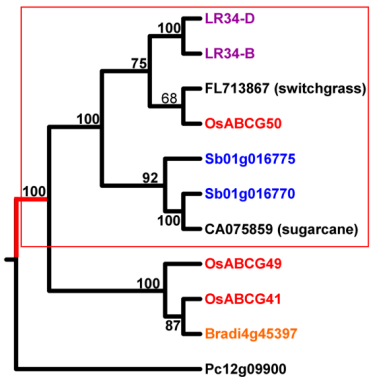


Figure 3. Phylogenetic tree of homologous LR34-proteins a) Phylogenetic tree including between 9 and 12 of the most homologous ABCG proteins from rice (red), *Brachypodium distachyon* (orange), sorghum (dark blue), and maize (light blue)**.** Furthermore, sequences of LR34-B and LR34-D (purple) as well as the two closest Arabidopsis homologs (green) and AtABCG36 (Pen3) were included. **b)** Tree including two EST sequences of switchgrass and sugarcane. The red line branch leads to the cluster that contained LR34-B and LR34-D. Orthologs are indicated by a red box. An ABC-like sequence from *Penicillium chrysogenum* (Pc12g09900) was used as outgroup to root the tree. Bootstrap numbers at the forks indicate how many times the sequences to the right of the fork occurred in the same group out of 100 trees. Strong bootstrap values of at least 75 are shown in bold.

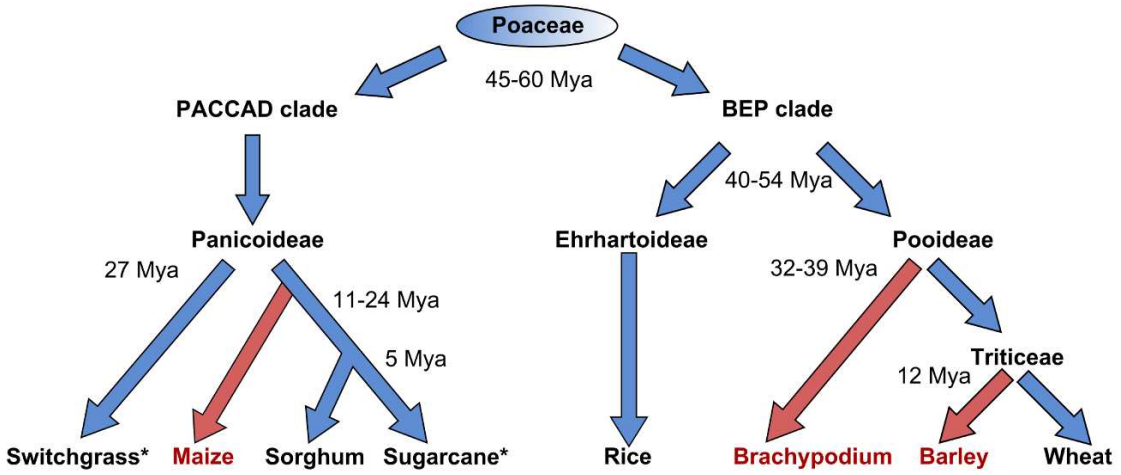


Figure 4. Distribution of the *Lr34* gene within Poaceae. The *Lr34* gene is represented in both clades of Poaceae (PACCAD; Panicoideae, Arundinoideae, Chlorideae, Centothecoideae, Aristidoideae and Danthonioideae) and BEP (Bambusoideae, Ehrhartoideae, Pooideae) suggesting it evolved before the separation of these two clades. *Lr34* has subsequently been lost in at least three lineages (maize, Brachypodium and barley – identified in red). The * denotes the presence of a *Lr34* EST. Predicted times of divergence are indicated and are according to Munkacsi *et al.* (2007), Chalupska *et al.* (2008), and The International Brachypodium Initiative (2010).

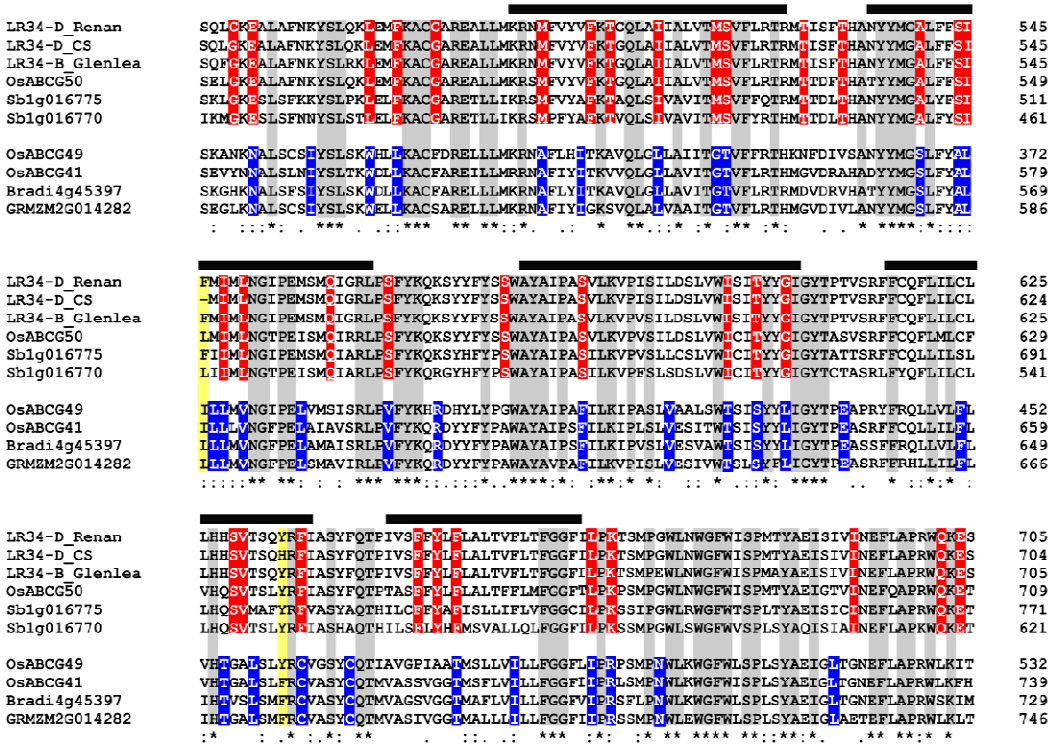


Figure 5. A snap-shot of the variation in amino acid sequence between LR34 orthologs and LR34-like ABCG proteins.

The sequence of the first transmembrane domain (with the first five (of six) transmembrane helices indicated with black bars) is shown. The prediction is based on TMAP (Persson and Argos 1997). The first six sequences represent LR34 orthologs and the remaining four comprise LR34-like proteins. Residues with 100% identity across both the LR34 orthologs and LR34-like families are highlighted in grey. Residues that are identical within the LR34 orthologs, but are not conserved at that position within the LR34-like proteins, are highlighted in red. Conversely, identical residues within the LR34-like ABCG proteins are highlighted in blue provided they are not represented in the LR34 orthologs. LR34sus-D (amino acids 466 to 705) from cultivar 'Renan' and LR34res-D from cultivar 'Chinese Spring (CS)' have been aligned with the homoeologous LR34-B copy of the resistant cultivar 'Glenlea' and the orthologs of rice and sorghum. The two amino acids that differed in LR34 between resistant and susceptible cultivars are highlighted in yellow. The LR34-like ABCG proteins are represented by two rice proteins, a protein from Brachypodium and one from maize. The complete sequences were aligned using ClustalX (Larkin *et al*. 2007). * represents identical amino acids, **:** represents highly conserved amino acids and **.** represents similar amino acids.

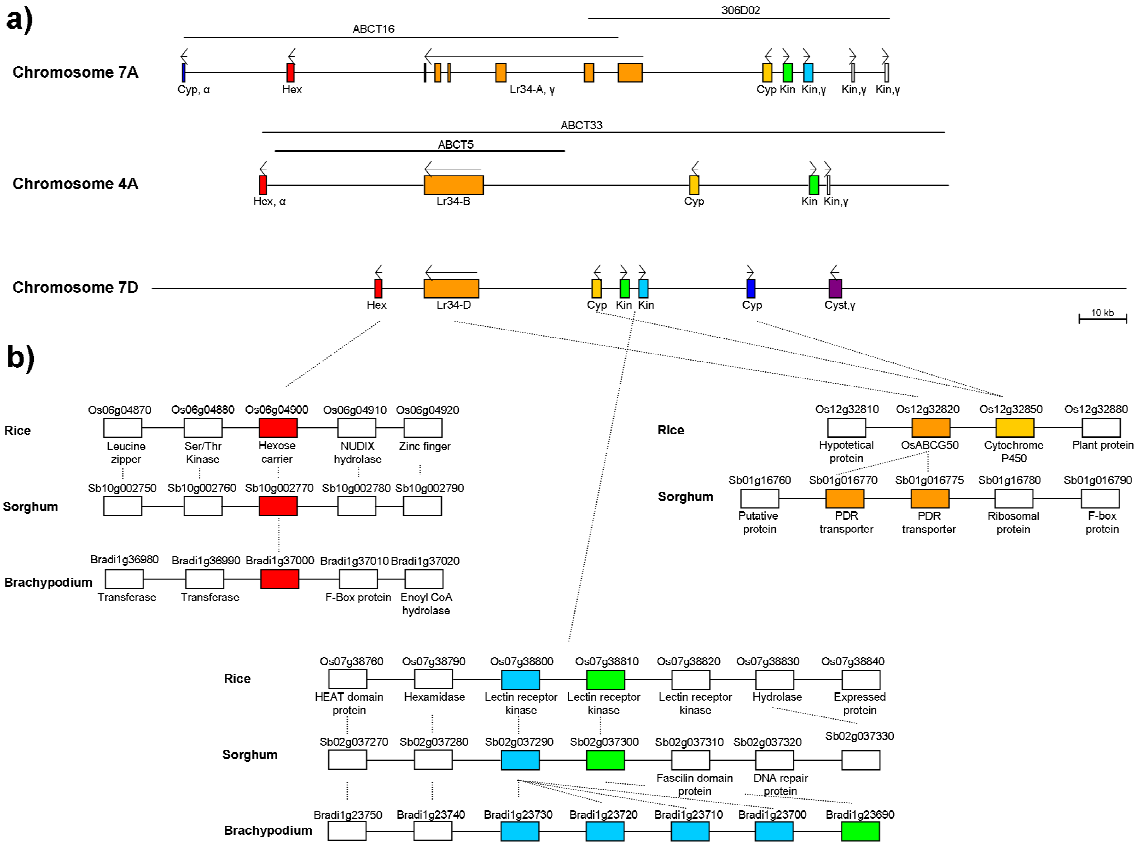


Figure 6. a) Genes present in the three homoeologous *Lr34* regions on chromosomes 7A, 4A, and 7D. Genes are represented as colored boxes. Hex = hexose carrier, Cyp = Cytochrome P450, Kin = lectin receptor kinase, Cyst = Cysteine proteinase. α = gene sequence is truncated because of BAC end, γ = pseudogene. **b)** Position of the orthologous genes found in the genome sequences of rice, sorghum, and *Brachypodium*. Rice genes Os12g32814 and Os12g32820 were re-annotated as one ABCG transporter gene.

Table S1. Meta QTL analysis including studies that identified QTLs in wheat for quantitative resistance against the three rusts of wheat and powdery mildew.

Population	Disease	Major QTLs	Minor QTLs	Reference
Arina x Forno	wheat stem rust	5BL, 7DS (Lr34, SWM10 - csLV34)	1AS, 7BL	Bansal *et al.* 2008
Saar x Avocet	wheat powdery mildew	1BL, 7DS (Lr34, gwm1220-SWM10), 4BL	3AS, 5AL, 5BS	Lillemo *et al.* 2008
Arina x Forno	wheat leaf rust	1BS (gwm604), 7DS (Lr34, cfd66-gwm1002)	2AL, 2DS, 2DL, 4BS, 7BS, 7BL	Schnurbusch *et al.* 2004
Fukuho-komugi x Oligoculm	wheat leaf rust, wheat stripe rust	1BL, 7DS (Lr34)	2DL, 3BS, 4BL, 4DL, 5BL, 6BS, 7BS, 7BL	Suenaga *et al.* 2003
Pavon 76 x Avocet	wheat leaf rust, wheat stripe rust	1BL, 6BL	3BS, 4BL, 6AL	William *et al.* 2006
Libellula x Huixianhong, Strampelli x Huixianhong	wheat stripe rust	2DS, 7DS (Lr34, csLV34-gwm295)	4BL, 5BL	Lu *et al.* 2009
Louise x Penawawa	wheat stripe rust	2BS		Carter *et al.* 2009
Beaver x Soissons	wheat leaf rust	1B, 4B, 5AS	1A, 3BS, 3BL, 3D, **4A (wtP-6447)**, 4D	Singh *et al.* 2009
TA4152-60 x ND495	wheat leaf rust	3AL, 3BL	4DL, 5BL, 6BL	Chu *et al.* 2009
Ning7840 x Chokwang	wheat leaf rust	1BL		Li *et al.* 2009
Express x Avocet S	wheat stripe rust	1BL, 3BL, 6AS		Lin *et al.* 2009
Guardian x Avocet S	wheat stripe rust	1BL	2D, 4B	Melicher *et al.* 2008
Aquileja x Luke	wheat stripe rust	2BS		Guo *et al.* 2008
Fukuho-komugi x Oligoculm	wheat powdery mildew	1AS, 7DS (Lr34)	2BL, 4BL	Liang *et al.* 2006
CI 13227 x Suwon 92	wheat leaf rust	2DS	2B, 7BL	Xu *et al.* 2005
Otane x Tiritea	wheat stripe rust	7DS (Lr34), 7BL	5DL	Imtiaz *et al.* 2004
Kariega x Avocet S	wheat stripe rust	2B, 7D (Lr34, gwm 295)	1A, **4A**, 7A	Ramburan *et al.* 2004
Opata 85 x W-7984	wheat stripe rust	2B, 7D (Lr34)	5A, 3D, 6D	Boukhatem *et al.* 2002
Forno x Oberkulmer	wheat leaf rust	7B	1BS, 2B, 3A, 4B, 4DL, 5DL	Messmer *et al.* 2000
Opata 85 x W-7984	leaf rust	7DS (Lr34), **4AL,** 2BS, 7BL		Faris *et al.* 1999
Parula x Siete cerros	leaf rust	1B, 1D, 7BL		William *et al.* 1997

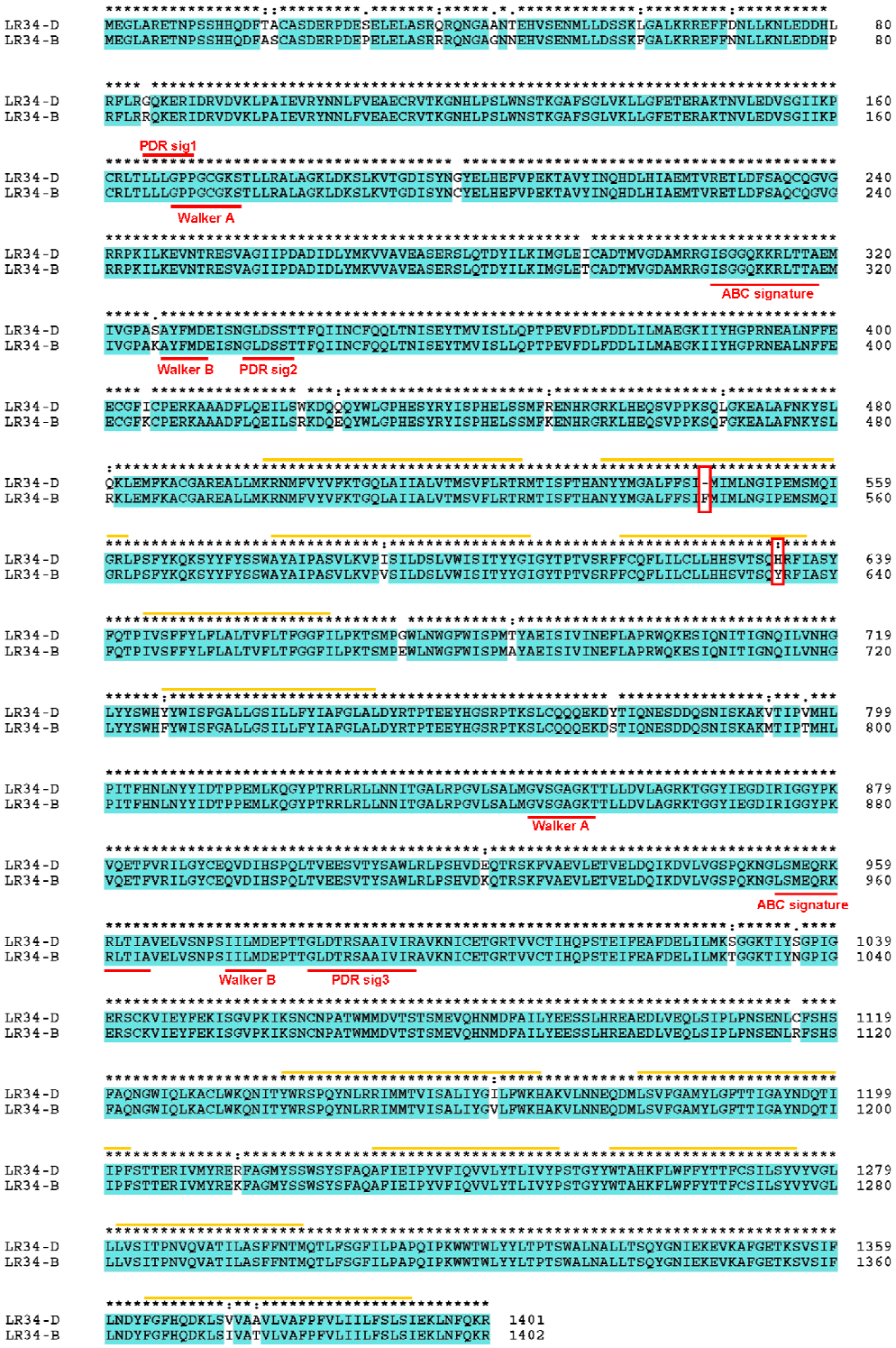


Figure S1. Alignment of LR34res-D from cultivar 'Chinese Spring' and LR34-B from 'Glenlea'. Conserved motifs of the two nucleotide binding domains are indicated as red lines. Yellow lines indicated transmembrane helices as predicted by TMAP. The two residues that differed between LR34res-D and LR34sus-D are indicated by red boxes.

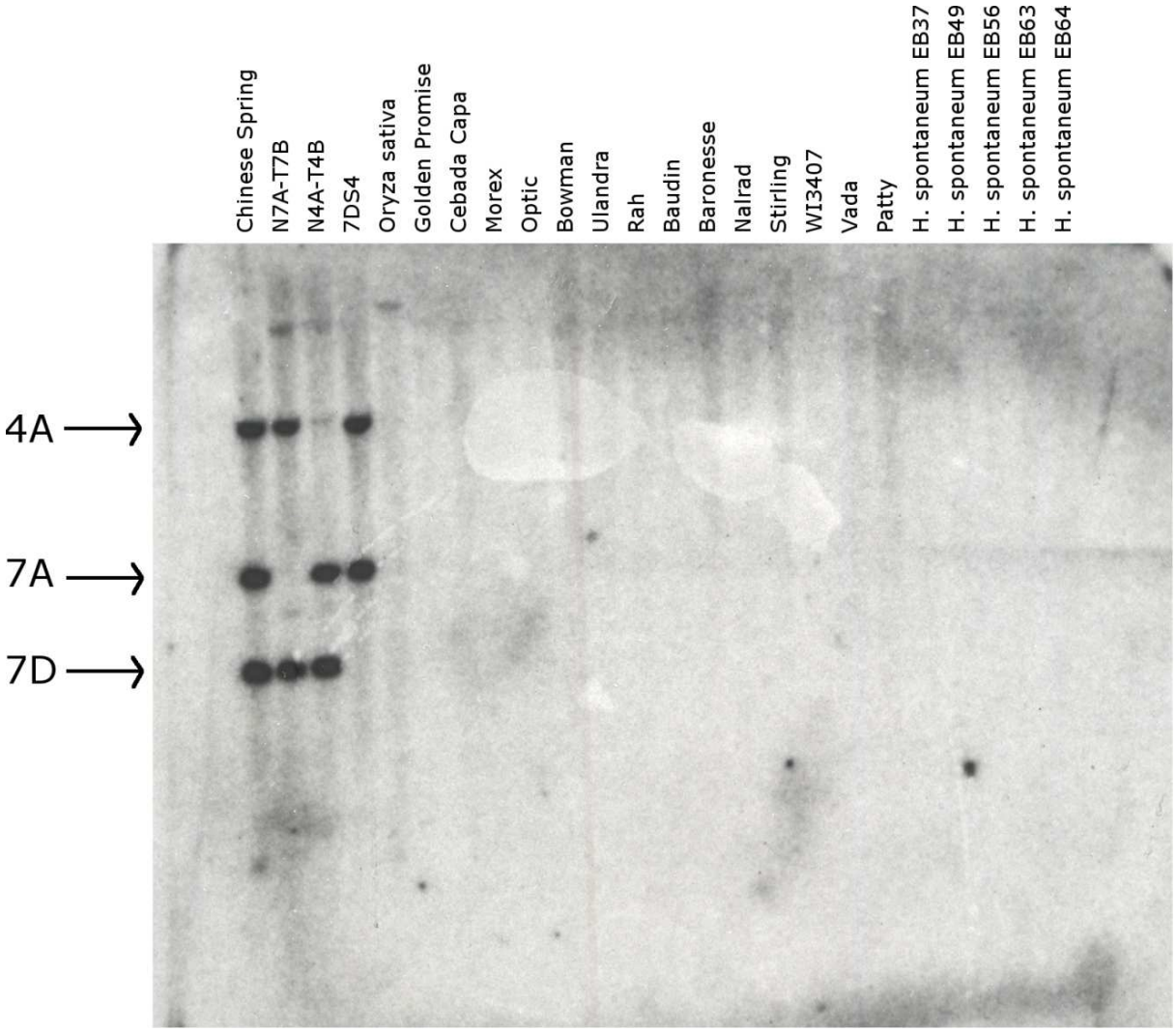


Figure S2. **Southern blot probed with a fragment of the *Lr34* gene.** The blot contains DNA of 14 barley cultivars and five wild barley accessions. Wheat line 'Chinese Spring', nulli-tetrasomic lines of homoeologous group 7 chromosomes, and rice were added as controls. Based on the gene sequences of *Lr34-A*, *Lr34-B*, *Lr34-D*, and OsABCG50 (Rice Genome Browser, Nipponbare subspecies of rice) we calculated the expected fragment sizes for the respective bands. They were 2,195 bp, 3,065 bp, 5,780 bp, and 12,254 bp for *Lr34-D*, *Lr34-A*, *Lr34-B*, and OsABCG50, respectively. Because of the relatively big size of the rice fragment, there is only a weak band at the very top of the respective lane.
N7A-T7B = nullisomic 7A-tetrasomic 7B; N4A-T4B = nullisomic 4A-tetrasomic 4B; 7DS4 = deletion line 7DS4 that lacks *Lr34*. Nulli-tetrasomic and deletion lines were derived from the Chinese landrace 'Chinese Spring'.

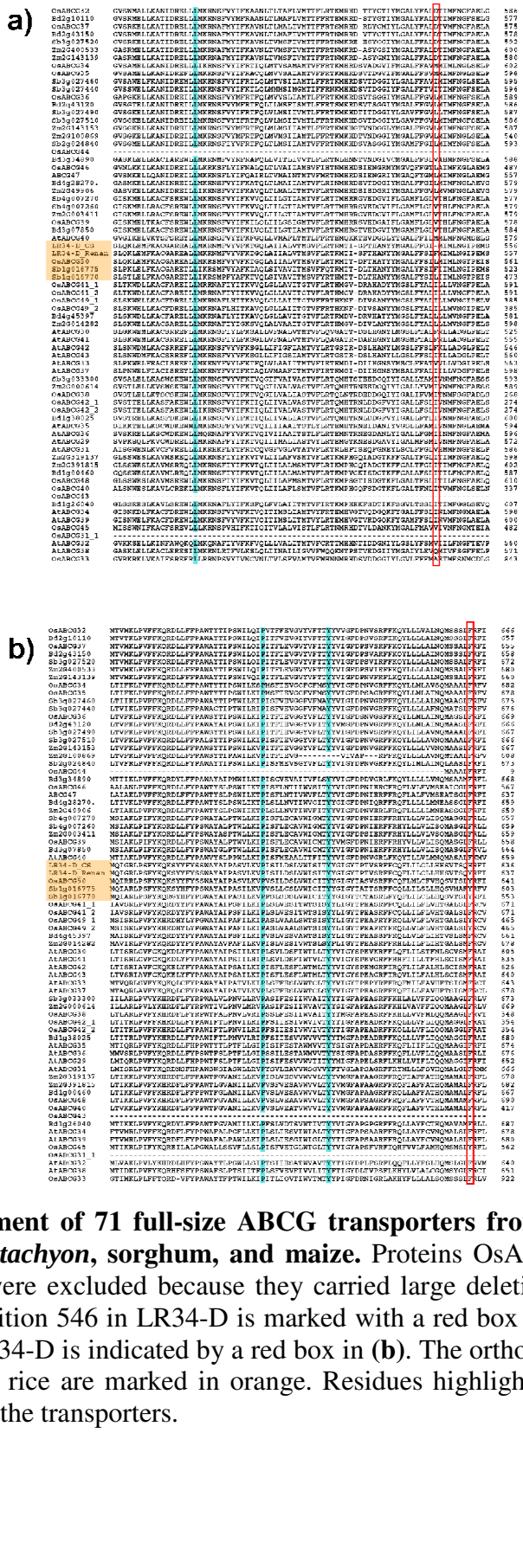


Figure S3: **Alignment of 71 full-size ABCG transporters from rice, Arabidopsis, *Brachypodium distachyon*, sorghum, and maize.** Proteins OsABCG31, OsABCG43, and OsABCG44 were excluded because they carried large deletions. The residue that corresponds to position 546 in LR34-D is marked with a red box in **(a)** and amino acid position 634 of LR34-D is indicated by a red box in **(b)**. The orthologous LR34 proteins from sorghum and rice are marked in orange. Residues highlighted in blue are 100% identical across all the transporters.

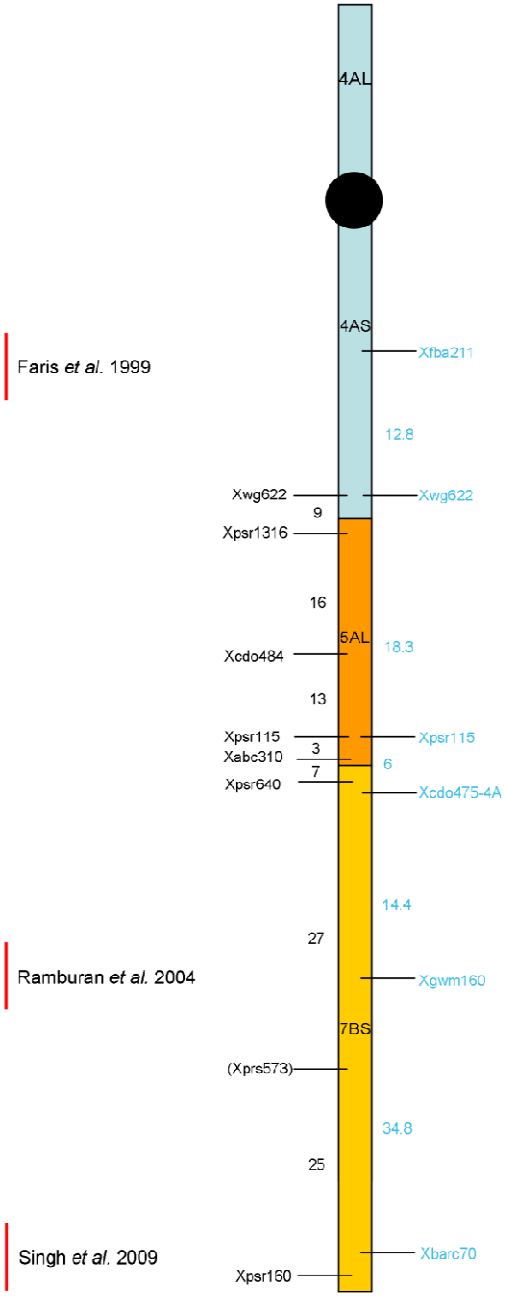


Figure S4. **Schematic representation of chromosome 4AL.** Several rearrangements occurred during the evolution of hexaploid wheat; a 4A/5AL translocation (orange), a pericentric inversion (4AS now corresponds to 4AL), and a 4AL/7BS translocation (yellow). Loci defining the translocation breakpoints are indicated in black (Devos *et al.* 1995). Markers found in the SSR map of Song *et al.* (2005) are indicated in blue font. For simplicity not all markers mapped by Song *et al.* (2005) are shown. The RFLP probe cdo475 was mapped to chromosomes 7A and 7D (Paillard *et al.* 2004, Schnurbusch *et al.* 2004) and to chromosome 4A (Song et al 2005). *Lr34-A* and *Lr34-D* were found to be closely linked to *Xcdo475-7A* and *Xcdo475-7D* (Schnurbusch *et al.* 2004), respectively. We therefore expect *Lr34-B* to map close to locus *Xcdo475-4A*. Red lines indicate the position of QTLs found to map on chromosome 4AL.

Akhunov, E.D., Akhunova, A.R., Linkiewicz, A.M., Dubcovsky, J., Hummel, D., Lazo, G., Chao, S., Anderson, O.D., David, J., Qi, L. *et al.* (2003) Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates. *Proc Natl Acad Sci U S A*, **100**, 10836-10841.

Allouis, S., Moore, G., Bellec, A., Sharp, R., Faivre Rampant, P., Mortimer, K., Pateyron, S., Foote, T.N., Griffiths, S., Caboche, M. *et al.* (2003) Construction and characterisation of a hexaploid wheat (*Triticum aestivum* L.) BAC library from the reference germplasm 'Chinese Spring'. *Cereal Research Communications*, **31**, 331-338.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J.H., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, **25**, 3389-3402.

Bossolini, E., Wicker, T., Knobel, P.A. and Keller, B. (2007) Comparison of orthologous loci from small grass genomes Brachypodium and rice: implications for wheat genomics and grass genome annotation. *Plant J*, **49**, 704-717.

Bottley, A., Xia, G.M. and Koebner, R.M. (2006) Homoeologous gene silencing in hexaploid wheat. *Plant J*, **47**, 897-906.

Buschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., vanDaelen, R., vanderLee, T., Diergaarde, P., Groenendijk, J. *et al.* (1997) The barley mlo gene: A novel control element of plant pathogen resistance. *Cell*, **88**, 695-705.

Chalupska, D., Lee, H.Y., Faris, J.D., Evrard, A., Chalhoub, B., Haselkorn, R. and Gornicki, P. (2008) Acc homoeoloci and the evolution of wheat genomes. *Proc Natl Acad Sci of the U S A*, **105**, 9691-9696.

Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Dossat, C., Sourdille, P., Joudrier, P. *et al.* (2005) Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *Plant Cell*, **17**, 1033-1045.

Crouzet, J., Trombik, T., Fraysse, A.S. and Boutry, M. (2006) Organization and function of the plant pleiotropic drug resistance ABC transporter family. *Febs Lett*, **580**, 1123-1130.

Dakouri, A., McCallum, B.D., Walichnowski, A.Z. and Cloutier, S. (2010) Fine-mapping of the leaf rust Lr34 locus in *Triticum aestivum* (L.) and characterization of large germplasm collections support the ABC transporter as essential for gene function. *Theor Appl Genet*, **121**, 373-384.

Devos, K.M., Dubcovsky, J., Dvorak, J., Chinoy, C.N. and Gale, M.D. (1995) Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. *Theor Appl Genet*, **91**, 282-288.

Faris, J.D., Li, W.L., Liu, D.J., Chen, P.D. and Gill, B.S. (1999) Candidate gene analysis of quantitative disease resistance in wheat. *Theor Appl Genet*, **98**, 219-225.

Fu, D., Uauy, C., Distelfeld, A., Blechl, A., Epstein, L., Chen, X., Sela, H., Fahima, T. and Dubcovsky, J. (2009) A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. *Science*, **323**, 1357-1360.

Fukuoka, S., Saka, N., Koga, H., Ono, K., Shimizu, T., Ebana, K., Hayashi, N., Takahashi, A., Hirochika, H., Okuno, K. *et al.* (2009) Loss of function of a

proline-containing protein confers durable disease resistance in rice. *Science*, **325**, 998-1001.

Gu, Y.Q., Salse, J., Coleman-Derr, D., Dupin, A., Crossman, C., Lazo, G.R., Huo, N.X., Belcram, H., Ravel, C., Charmet, G. *et al.* (2006) Types and rates of sequence evolution at the high-molecular-weight glutenin locus in hexaploid wheat and its ancestral genomes. *Genetics*, **174**, 1493-1504.

Hirokawa, T., Boon-Chieng, S. and Mitaku, S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics*, **14**, 378-379.

Hoisington, D., Khairallah, M., Reeves, T., Ribaut, J., Skovmand, B., Taba, S. and Warburton, M. (1999) Plant genetic resources: What can they contribute toward increased crop productivity? *Proc Natl Acad Sci U S A,* **96**, 5937-5943.

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature*, **436**, 793-800.

Jasinski, M., Ducos, E., Martinoia, E. and Boutry, M. (2003) The ATP-binding cassette transporters: Structure, function, and gene family comparison between rice and Arabidopsis. *Plant Physiology*, **131**, 1169-1177.

Keller, B., Feuillet, C. and Yahiaoui, N. (2005) Map-based isolation of disease resistance genes from bread wheat: cloning in a supersize genome. *Genet Res*, **85**, 93-100.

Kolmer, J.A. (2005) Tracking wheat rust on a continental scale. *Curr Opin Plant Biol,* **8,** 441-449**.**

Kolmer, J.A., Singh, R.P., Garvin, D.F., Viccars, L., William, H.M., Huerta-Espino, J., Ogbonnaya, F.C., Raman, H., Orford, S., Bariana, H.S. *et al.* (2008) Analysis of the Lr34/Yr18 rust resistance region in wheat germplasm. *Crop Sci*, **48**, 1841-1852.

Kou, Y.J. and Wang, S.P. (2010) Broad-spectrum and durability: understanding of quantitative disease resistance. *Curr Opin Plant Biol*, **13**, 181-185.

Krattinger, S.G., Lagudah, E.S., Spielmeyer, W., Singh, R.P., Huerta-Espino, J., McFadden, H., Bossolini, E., Selter, L.L. and Keller, B. (2009) A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science*, **323**, 1360-1363.

Lagudah, E.S., Krattinger, S.G., Herrera-Foessel, S., Singh, R.P., Huerta-Espino, J., Spielmeyer, W., Brown-Guedira, G., Selter, L.L. and Keller B. (2009) Gene-specific markers for the wheat gene *Lr34/Yr18/Pm38* which confers resistance to multiple fungal pathogens. *Theor Appl Genet,* **119**, 889-898.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947-2948.

Mickelson-Young, L., Endo, T.R. and Gill, B.S. (1995) A cytogenetic ladder-map of the wheat homoeologous group-4 chromosomes. *Theor Appl Genet*, **90**, 1007-1011.

Munkacsi, A.B., Stoxen, S. and May, G. (2007) Domestication of maize, sorghum, and sugarcane did not drive the divergence of their smut pathogens. *Evolution*, **61**, 388-403.

Naranjo, T., Roca, A., Goicoechea, P.G. and Giraldez, R. (1987) Arm homoeology of wheat and rye chromosomes. *Genome*, **29**, 873-882.

Nilmalgoda, S.D., Cloutier, S. and Walichnowski, A.Z. (2003) Construction and characterization of a bacterial artificial chromosome (BAC) library of hexaploid wheat (*Triticum aestivum* L.) and validation of genome coverage using locus-specific primers. *Genome*, **46**, 870-878.

Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., Thibaud-Nissen, F., Malek, R.L., Lee, Y., Zheng, L. *et al.* (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res*, **35**, D883-887.

Paillard, S., Schnurbusch, T., Winzeler, M., Messmer, M., Sourdille, P., Abderhalden, O., Keller, B. and Schachermayr, G. (2003) An integrative genetic linkage map of winter wheat (*Triticum aestivum* L.). *Theor Appl Genet*, **107**, 1235-1242.

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A. *et al.* (2009) The Sorghum bicolor genome and the diversification of grasses. *Nature*, **457**, 551-556.

Persson, B. and Argos, P. (1997) Prediction of membrane protein topology utilizing multiple sequence alignments. *J Protein Chem*. **16**, 453-457.

Poland, J.A., Balint-Kurti, P.J., Wisser, R.J., Pratt, R.C. and Nelson, R.J. (2009) Shades of gray: the world of quantitative disease resistance. *Trends Plant Sci*, **14**, 21-29.

Pumphrey, M., Bai, J., Laudencia-Chingcuanco, D., Anderson, O. and Gill, B.S. (2009) Nonadditive expression of homoeologous genes is established upon polyploidization in hexaploid wheat. *Genetics*, **181**, 1147-1157.

Ramburan, V.P., Pretorius, Z.A., Louw, J.H., Boyd, L.A., Smith, P.H., Boshoff, W.H.P. and Prins, R. (2004) A genetic analysis of adult plant resistance to stripe rust in the wheat cultivar Kariega. *Theor Appl Genet*, **108**, 1426-1433.

Rea, P.A. (2007) Plant ATP-Binding cassette transporters. *Annu Rev Plant Biol*, **58**, 347-375.

Roder, M.S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M.H., Leroy, P. and Ganal, M.W. (1998) A microsatellite map of wheat. *Genetics*, **149**, 2007-2023.

Sambrook, J. and Russel, D.W. (2001) *Molecular Cloning: a Laboratory Manual* NY, USA: Cold Spring Harbor Laboratory Press.

Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T.A. *et al.* (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science*, **326**, 1112-1115.

Schnurbusch, T., Bossolini, E., Messmer, M. and Keller, B. (2004) Tagging and validation of a major quantitative trait locus for leaf rust resistance and leaf tip necrosis in winter wheat cultivar forno. *Phytopathology*, **94**, 1036-1041.

Sears, S.R. (1954) The aneuploids of common wheat. *Mo Agric Exp Stn Res Bull*, **572**, 1-58.

Singh, D., Simmonds, J., Park, R.F., Bariana, H.S. and Snape, J.W. (2009) Inheritance and QTL mapping of leaf rust resistance in the European winter wheat cultivar 'Beaver'. *Euphytica*, **169**, 253-261.

Song, Q.J., Shi, J.R., Singh, S., Fickus, E.W., Costa, J.M., Lewis, J., Gill, B.S., Ward, R. and Cregan, P.B. (2005) Development and mapping of microsatellite (SSR) markers in wheat. *Theor Appl Genet*, **110**, 550-560.

Sorrells, M.E., La Rota, M., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., Munkvold, J.D., Miftahudin, Mahmoud, A., Ma, X.F., Gustafson, P.J. *et al.* (2003) Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res*, **13**, 1818-1827.

Stein, M., Dittgen, J., Sanchez-Rodriguez, C., Hou, B.H., Molina, A., Schulze-Lefert, P., Lipka, V. and Somerville, S. (2006) Arabidopsis PEN3/PDR8, an ATP binding cassette transporter, contributes to nonhost resistance to inappropriate pathogens that enter by direct penetration. *Plant Cell*, **18**, 731-746.

Strange, R.N. and Scott, P.R. (2005) Plant disease: a threat to global food security. *Annu Rev Phytopathol*, **43**, 83-116.

The International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass Brachypodium distachyon. *Nature*, **463**, 763-768.

Udall, J.A. and Wendel, J.F. (2006) Polyploidy and crop improvement. *Crop Sci*, **46**, S3-S14.

Van Deynze, A.E., Sorrells, M.E., Park, W.D., Ayres, N.M., Fu, H., Cartinhour, S.W., Paul, E. and McCouch, S.R. (1998) Anchor probes for comparative mapping of grass genera. *Theor Appl Genet*, **97**, 356-369.

Wicker, T., Krattinger, S.G., Lagudah, E.S., Komatsuda, T., Pourkheirandish, M., Matsumoto, T., Cloutier, S., Reiser, L., Kanamori, H., Sato, K. *et al.* (2009) Analysis of intraspecies diversity in wheat and barley genomes identifies breakpoints of ancient haplotypes and provides insight into the structure of diploid and hexaploid triticeae gene pools. *Plant Physiol*, **149**, 258-270.

Wicker, T., Matthews, D.E. and Keller, B. (2002) TREP: a database for Triticeae repetitive elements. *Trends in Plant Science*, **7**, 561-562.

Woodhouse, M.R., Schnable, J.C., Pedersen, B.S., Lyons, E., Lisch, D., Subramaniam, S. and Freeling, M. Following Tetraploidy in Maize, a Short Deletion mechanism removed genes preferentially from one of the two homeologs. *PLOS Biology*, e10000409.